

CLUSTERING ASSIGNMENT

Clustering Countries Based On Socio-economic For Financial Aid

**Submitted By:
Abhishek De**

BACKGROUND

HELP International is an international humanitarian NGO that is committed to fighting poverty and providing the people of backward countries with basic amenities and relief during the time of disasters and natural calamities. It runs a lot of operational projects from time to time along with advocacy drives to raise awareness as well as for funding purposes.

After the recent funding program's, they have been able to raise around \$ 10 million. Now the CEO of the NGO needs to decide how to use this money strategically and effectively. The significant issues that come while making this decision are mostly related to choosing the countries that are in the direst need of aid.

BUSINESS OBJECTIVE

Our job is to categorize the countries using some socio-economic and health factors that determine the overall development of the country. Then we need to suggest the countries which the CEO needs to focus on the most.

TECHNICAL APPROACH

- Using Hierarchical Clustering to identify the optimal cluster value.
- Use Silhouette and Elbow method to validate the optimal cluster value.
- Use K-Means Cluster method to build the final cluster model.
- Identify appropriate cluster for Financial Aid using Cluster Mean method.
- Analyze the Final Cluster Statistics against other clusters.
- Decision making on the final list based on the Descriptive Statistics of our Final Cluster.
- Choose the Top-10 Countries from the Final Cluster based on the Higher Child Mortality, Lower Income and Low GDP.
- Present the Final Report.

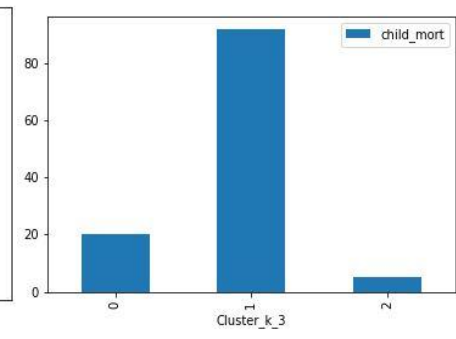
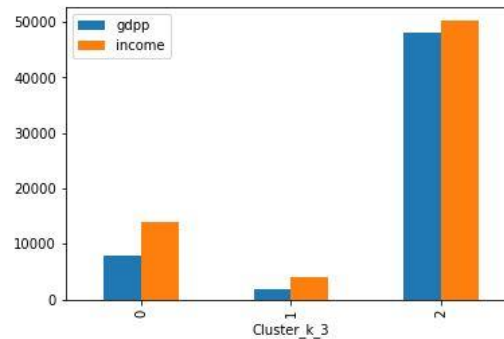
CLUSTER SUMMARY

Cluster '**Under-Developed Countries**' has the highest average **Child Mortality** rate of **~92**, when compared to other 3 clusters and average **GDPP** and **Income** of **~1909** & **~3897** respectively.

All these figures clearly make this cluster the **best candidate** for the **Financial Aid** from NGO.

We could see that cluster '**Under-Developed Countries**' comprises of 29% of over-all data and has 48 obs. In comparison to 167 total obs.

Cluster Name	child_mort	income	gdpp	Obs.	Prop.
Under-Developed Countries	91.61	3897.35	1909.21	48.00	0.20
Developing Countries	20.36	13968.02	7979.91	91.00	0.54
Developed Countries	5.05	50178.57	48114.29	28.00	0.17



Cluster Division of the Countries

FINAL LIST OF COUNTRIES FROM 'UNDER-DEVELOPED COUNTRIES'

We concluded on the top 10 countries from the final cluster ('Under Developed Country') based on the cluster median values of Child Mortality, Income and GDPP.

We filtered the countries with:

- **Lowest GDP.**
- **Lowest Income**
- **Highest Child Mortality**

Country
Burundi
Congo, Dem. Rep.
Central African Republic
Malawi
Guinea-Bissau
Afghanistan
Burkina Faso
Guinea
Haiti
Benin

Top 10 countries for Financial Aid.

STATISTICS

	income	child_mort	gdpp
count	10.000000	10.000000	10.000000
mean	1223.100000	119.730000	521.300000
std	393.260572	35.448744	158.230528
min	609.000000	90.200000	231.000000
25%	923.500000	97.450000	449.250000
50%	1290.000000	112.500000	550.000000
75%	1482.500000	116.000000	629.750000
max	1820.000000	208.000000	758.000000

Descriptive Statistics

	country	income	child_mort	gdpp
0	Afghanistan	1610	90.2	553
17	Benin	1820	111.0	758
25	Burkina Faso	1430	116.0	575
26	Burundi	764	93.6	231
31	Central African Republic	888	149.0	446
37	Congo, Dem. Rep.	609	116.0	334
63	Guinea	1190	109.0	648
64	Guinea-Bissau	1390	114.0	547
66	Haiti	1500	208.0	662
94	Malawi	1030	90.5	459

Country-wise Statistics

DISTRIBUTION OF ATTRIBUTES

